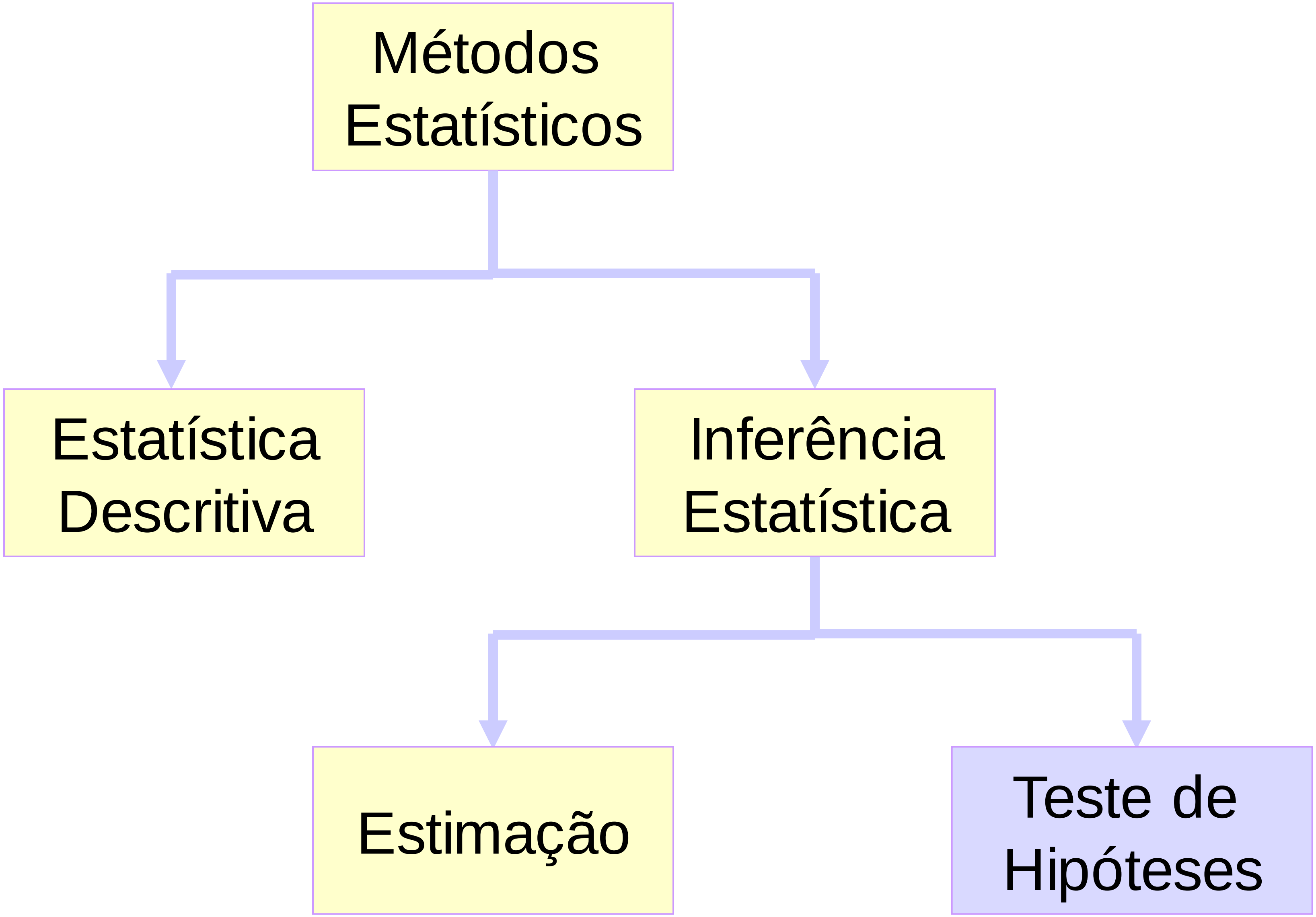
Teste de hipóteses para proporção populacional p

Métodos Estatísticos



TESTE DE HIPÓTESES

Estimação

Qual é a probabilidade de "cara" no lançamento de uma moeda?

Qual é a taxa média de glicose em mulheres com mais de 60 anos?

Qual é a proporção de moradores do RJ, com idades entre 15 e 50 anos, que contraíram a dengue em 2013?

Teste de Hipóteses

A moeda é honesta ou é desequilibrada?

A taxa média de glicose em mulheres com mais de 60 anos é superior a 100 mg/ml?

Pelo menos 2% dos moradores do RJ, com idades entre 15 e 50 anos, contraíram a dengue em 2013?

Introdução

Em **estimação** o objetivo é “estimar” o valor desconhecido de um parâmetro, por exemplo, da proporção $\boldsymbol{p}$ de “indivíduos” em uma população com determinada característica ou da média $\boldsymbol{\mu}$ de uma variável X.

A estimativa é baseada em uma amostra casual simples de tamanho $\boldsymbol{n}$.

Entretanto, se o objetivo for saber se a estimativa pontual observada na amostra dá ou não suporte a uma conjectura sobre o valor de parâmetro, trata-se de **testar hipóteses.**

O que é uma hipótese?

- É uma conjectura sobre um <u>parâmetro populacional</u>.

  Por exemplo, a proporção p é um parâmetro populacional.

- A hipótese deve ser estabelecida antes da análise.

Exemplo 1: Queremos avaliar se uma moeda é honesta.

Ou seja, queremos testar a

hipótese nula H_0: a moeda é honesta

contra a

hipótese alternativa H_1: a moeda não é honesta

Em linguagem estatística, essas hipóteses podem ser reescritas como:

$$H_0: p = 0{,}5$$

$$H_1: p \neq 0{,}5$$

com p sendo a probabilidade de "cara" da moeda.

Obs.: Nesse caso, dizemos que a hipótese alternativa é **bilateral.**

Hipóteses

⇒ Como estabelecer as hipóteses estatísticas do teste?

No caso especial de teste de hipóteses sobre o parâmetro p, temos:

Hipótese nula: afirmação sobre p, em geral, ligada a um valor de referência, ou a uma especificação padrão ou histórica.

Hipótese alternativa: afirmação sobre p que suspeitamos seja verdadeira.

Se observarmos 30 caras em 50 lançamentos independentes da moeda, implicando $\hat{p} = 0{,}60$, o que podemos concluir?

E se observarmos 20 caras $(\hat{p} = 0{,}40)$?

ou 10 caras $(\hat{p} = 0{,}20)$? ou 45 caras $(\hat{p} = 0{,}90)$?

Podemos considerar uma **regra de decisão**, como por exemplo,

"Se, em 50 lançamentos da moeda, observarmos

$$\hat{p} \leq 0{,}35 \text{ ou } \hat{p} \geq 0{,}65$$

então, rejeitamos a hipótese nula H_0 de que a moeda seja honesta; caso contrário, não rejeitamos a hipótese H_0."

Testar uma hipótese estatística é estabelecer uma **regra** que nos permita, com base na informação de uma amostra, **decidir pela rejeição ou não de H_0.**

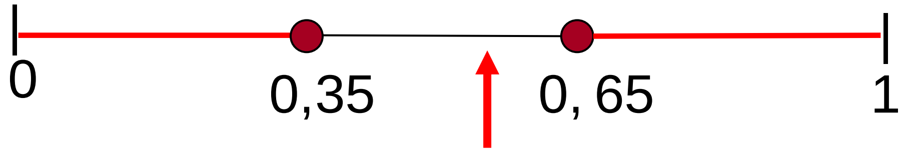
No exemplo, segundo a regra de decisão, o conjunto de valores de $\hat{p}$ que levam à rejeição da hipótese nula H_0 é $\{\hat{p} : \hat{p} \leq 0{,}35$ ou $\hat{p} \geq 0{,}65\}$, o qual denominamos de **região crítica (*RC*)** ou **região de rejeição de H_0**, ou seja,

$RC = \{\hat{p} : \hat{p} \leq 0{,}35$ ou $\hat{p} \geq 0{,}65\}$: **região de rejeição**

$RC^c = \{\hat{p} : 0{,}35 < \hat{p} < 0{,}65\}$: **região de não rejeição de H_0**

Regra de decisão (teste)

No exemplo da moeda, suponha que observemos 30 caras, isto é, $\hat{p} = 0{,}6.$



Valor observado na amostra

$\hat{p} \notin RC$ $\Rightarrow$ **Não rejeitamos H_0**

Agora suponha que observemos 10 caras, isto é, $\hat{p} = 0{,}20.$

Valor observado na amostra

$\hat{p} \in RC$ $\Rightarrow$ **Rejeitamos H_0**

Regra de decisão (teste):

$$\hat{p} \in RC \Rightarrow \textbf{ rejeitamos } H_0$$

$$\hat{p} \notin RC \Rightarrow \textbf{ não rejeitamos } H_0$$

Será que nossa conclusão está correta?

Ao decidir pela rejeição ou não da hipótese nula H_0, podemos cometer *dois tipos de erro*.

Erros

Erro tipo I: Rejeitar H_0 quando H_0 é verdadeira

(afirmar que a moeda não é honesta quando, na verdade, ela é).

Erro tipo II: Não rejeitar H_0 quando H_0 é falsa

(afirmar que a moeda é honesta quando, na verdade, ela é desequilibrada).

Exemplo: Um medicamento genérico é lançado no mercado.

As hipóteses são:

H_0: *O medicamento é inequivalente ao medicamento de marca*.

H_1: O medicamento é quivalente ao medicamento de marca.

- **Erro I**: O medicamento é equivalente ao medicamento de marca apesar de ser inequivalente.
- **Erro II**: O medicamento é inequivalente ao medicamento de marca apesar de ser equivalente.

Naturalmente, deve-se reduzir a possibilidade de ocorrer o Erro I, pois entende-se que é mais grave lançar um medicamento genérico ineficaz do que proibir um medicamento genérico eficaz.

Probabilidades de erros

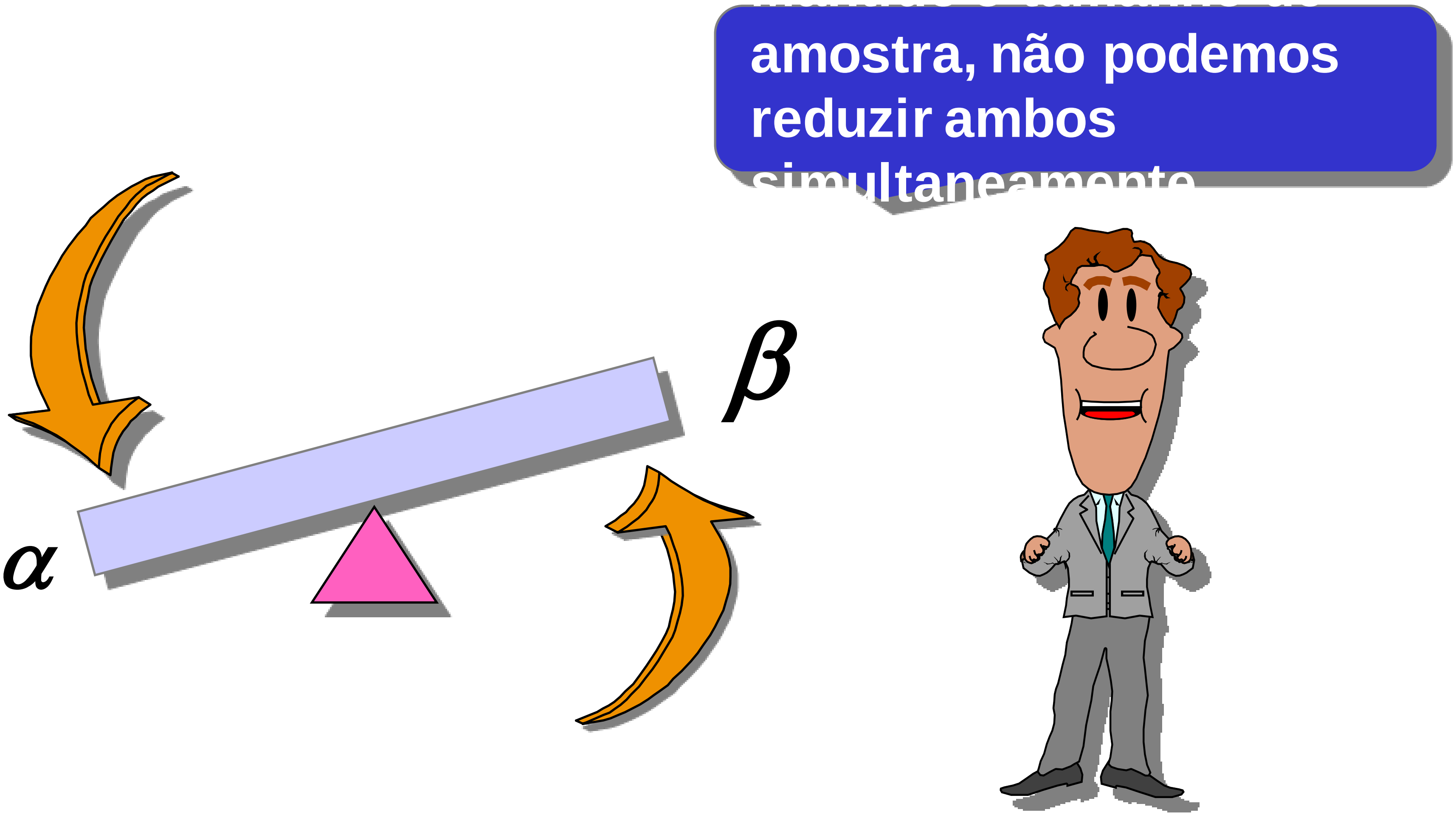
P**(erro I)** **=** P**(**rejeitar H_0 **|** H_0 é verdadeira**)** **=** α

α **: nível de significância do teste**

P**(erro II)** **=** P**(**não rejeitar H_0 **|** H_0 é falsa**)** **=** β

1 - β **: poder do teste**

- α e β tem uma relação inversa



- em geral, só podemos controlar um dos erros.

No exemplo da moeda,

$$H_0: \boldsymbol{p} = 0{,}5$$

$$H_1: \boldsymbol{p} \neq 0{,}5$$

$$RC = \{\hat{p} : \hat{p} \leq 0{,}35 \text{ ou } \hat{p} \geq 0{,}65\}$$

$$\alpha = P(\text{erro I}) = P(\text{rejeitar } H_0 \mid H_0 \text{ verdadeira })$$

$$= P(\hat{p} \in RC \mid p = 0{,}5) = P(\hat{p} \leq 0{,}35 \text{ ou } \hat{p} \geq 0{,}65 \mid p = 0{,}5)$$

$\Rightarrow$ Como calcular essa probabilidade?

Resultado 4: Teorema Limite Central (TLC)

Seja X uma $v.$ $a.$ que tem média μ e variância σ^2.

Para amostras X_1, X_2, ..., X_n , retiradas ao acaso e com reposição de X, a distribuição de probabilidade da média amostral $\overline{X}$ *aproxima-se*, *<u>para n grande</u>*, de uma distribuição normal, com média μ e variância σ^2/n , ou seja,

$$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right), \text{ para } n \text{ grande, aproximadamente.}$$

No caso da proporção, $X \sim Bernoulli(p)$. Então,

$$\hat{p} \sim N(p, \frac{p(1-p)}{n}), \text{ quando } n \text{ é grande}$$

Assim, sob H_0 (p = 0,5),

$$\hat{p} \sim N(0{,}5;\ \frac{0{,}5 \times 0{,}5}{50}) \text{, aprox.} \quad \Rightarrow Z = \frac{\hat{p} - 0{,}5}{\sqrt{0{,}25/50}} \sim N(0;1) \text{, aprox.}$$

Portanto, nesse caso,

$$\alpha = P(\hat{p} \leq 0{,}35 \text{ ou } \hat{p} \geq 0{,}65 \mid p = 0{,}5)$$

$$\cong P(Z \leq \frac{0{,}35 - 0{,}5}{\sqrt{0{,}25/50}}) + P(Z \geq \frac{0{,}65 - 0{,}5}{\sqrt{0{,}25/50}})$$

$$= P(Z \leq -2{,}12) + P(Z \geq 2{,}12) = 2 \times P(Z \geq 2{,}12)$$

$$= 2 \times (1 - 0{,}983) = 2 \times 0{,}017$$

$$= 0{,}034.$$

Se alterarmos a regra de decisão para

$$RC = \{\hat{p} : \hat{p} \leq 0{,}30 \text{ ou } \hat{p} \geq 0{,}70\}$$

o que acontece com o nível de significância do teste α (probabilidade de erro tipo I)?

Regiões críticas e níveis de significância α

RC	α
$\{\hat{p} : \hat{p} \leq 0{,}40 \text{ ou } \hat{p} \geq 0{,}60\}$	**0,1586**
$\{\hat{p} : \hat{p} \leq 0{,}35 \text{ ou } \hat{p} \geq 0{,}65\}$	**0,0340**
$\{\hat{p} : \hat{p} \leq 0{,}30 \text{ ou } \hat{p} \geq 0{,}70\}$	**0,0048**

$$RC \downarrow \qquad \alpha \downarrow \qquad \beta \uparrow$$

Considerando $RC = \{\hat{p} : \hat{p} \le 0{,}35 \text{ ou } \hat{p} \ge 0{,}65\}$

Decisão	**Verdadeiro valor de p**	
	p **= 0,5** (H_0 é verd.)	$p \neq$ **0,5** (H_1 é verd.)
Não rejeitar H_0	Decisão correta **1 -** α = 0,966	**Erro II** β
Rejeitar H_0	**Erro I** α = 0,034	Decisão correta **1 -** β

Até agora, o procedimento foi

escolher $RC \Rightarrow$ **determinar** α

Alternativamente, podemos

fixar $\alpha \Rightarrow$ **determinar** RC

Os valores de nível de significância α, usualmente adotados, são entre 1% e 10%.

Determinação da região crítica

Exemplo 2: Suponha que um medicamento existente no mercado produza o efeito desejado em 60% dos casos nos quais é aplicado.

Um laboratório produz um **novo medicamento** e afirma que ele é melhor do que o existente.

Objetivo: Verificar, estatisticamente, se a afirmação do laboratório é verdadeira.

$\Rightarrow$ Aplicou-se o novo medicamento em n = **50** pacientes.

Seja $\boldsymbol{p}$ a probabilidade do novo medicamento ser eficaz ou a proporção populacional de pacientes para os quais o novo medicamento é eficaz.

(1) Hipóteses estatísticas:

$$\boxed{\begin{array}{l} H_0\text{: } \boldsymbol{p} = 0{,}6 \\ H_1\text{: } \boldsymbol{p} > 0{,}6 \end{array}}$$

que correspondem a

$\boldsymbol{H_0}$: o novo medicamento é similar ao existente

$\boldsymbol{H_1}$: o novo medicamento é melhor, mais eficaz

(2) Fixemos o nível de significância em 5% ($\alpha = 0{,}05$).

(3) A região crítica deve ter a forma:

$RC = \{\hat{p} \geq a\}$ $\Rightarrow$ Como obter o valor a?

O valor de $\boldsymbol{a}$ deve ser tal que

P(**erro I**) = $P(\hat{p} \in RC \mid p = 0{,}6) = P(\hat{p} \geq a \mid p = 0{,}6) = \alpha$

$$\Rightarrow \; 0{,}05 = P(\hat{p} \geq a \mid p = 0{,}6) \cong P\left(Z \geq \frac{a - 0{,}6}{\sqrt{0{,}24/50}}\right) \quad z$$

Pela tabela, para $A(z)$=0,95, temos z =1,64, ou seja,

$$\frac{a - 0{,}6}{\sqrt{0{,}24/50}} = 1{,}64 \Rightarrow a = 0{,}6 + 1{,}64\sqrt{0{,}24/50} \cong 0{,}714.$$

Portanto, $RC = \{\hat{p} \geq 0{,}714\}$.

Suponha que em 38 dos 50 pacientes o novo medicamento foi eficaz, ou seja, $\hat{p}_{obs} = 0{,}76.$

$\hat{p}_{obs} \in RC \Rightarrow H_0$ é rejeitada, isto é, concluímos ao nível de significância de 5 % que o novo medicamento é mais eficaz.

Resumo

(0) Definir o parâmetro $\boldsymbol{p}$ de interesse no problema.

(1) Estabelecer as **hipóteses estatísticas**:

H_0: $p = p_0$ contra uma das alternativas

H_1: $p \neq p_0$, H_1: $p > p_0$ ou H_1: $p < p_0$.

$\Downarrow$ $\Downarrow$ $\Downarrow$

bilateral **unilateral** **unilateral**

(2) Escolher um **nível de significância** α.

(3) Determinar a **região crítica** RC da forma

$$\{\hat{p} \leq a_1, \hat{p} \geq a_2\}, \ \{\hat{p} \geq a\}, \ \{\hat{p} \leq a\},$$

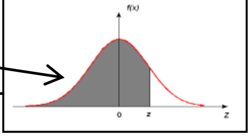
respectivamente às hipóteses alternativas.

(4) Selecionar uma **amostra** casual simples e determinar a proporção $\hat{p}$ de “indivíduos” na amostra portadores do atributo desejado.

(5) **Decidir**, usando a evidência $\hat{p}$, ao nível de significância α, e **concluir**.

$$\hat{p} \in RC \;\Rightarrow\; \textbf{rejeitamos } H_0$$
$$\hat{p} \notin RC \;\Rightarrow\; \textbf{não rejeitamos } H_0$$

Distribuição Normal : Valores de P(Z ≤ z) = A(z)



Parte inteira e primeira decimal de z

	Segunda decimal de z									
	0	1	2	3	4	5	6	7	8	9
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000